



## Fast ADM1 implementation for the optimization of feeding strategy using near infrared spectroscopy

Cyrille Charnier, Eric Latrille, Julie Jimenez, Michel Torrijos, Philippe Sousbie, Jérémie Miroux, Jean-Philippe Steyer

### ► To cite this version:

Cyrille Charnier, Eric Latrille, Julie Jimenez, Michel Torrijos, Philippe Sousbie, et al.. Fast ADM1 implementation for the optimization of feeding strategy using near infrared spectroscopy. Water Research, 2017, 122, pp.27-35. 10.1016/j.watres.2017.05.051 . hal-01603275

**HAL Id: hal-01603275**

**<https://hal.science/hal-01603275>**

Submitted on 26 May 2020

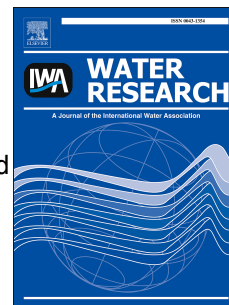
**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Accepted Manuscript

Fast ADM1 implementation for the optimization of feeding strategy using near infrared spectroscopy

Cyrille Charnier, Eric Latrille, Julie Jimenez, Michel Torrijos, Philippe Sousbie, Jérémie Miroux, Jean-Philippe Steyer



PII: S0043-1354(17)30371-8

DOI: [10.1016/j.watres.2017.05.051](https://doi.org/10.1016/j.watres.2017.05.051)

Reference: WR 12930

To appear in: *Water Research*

Received Date: 15 March 2017

Revised Date: 5 May 2017

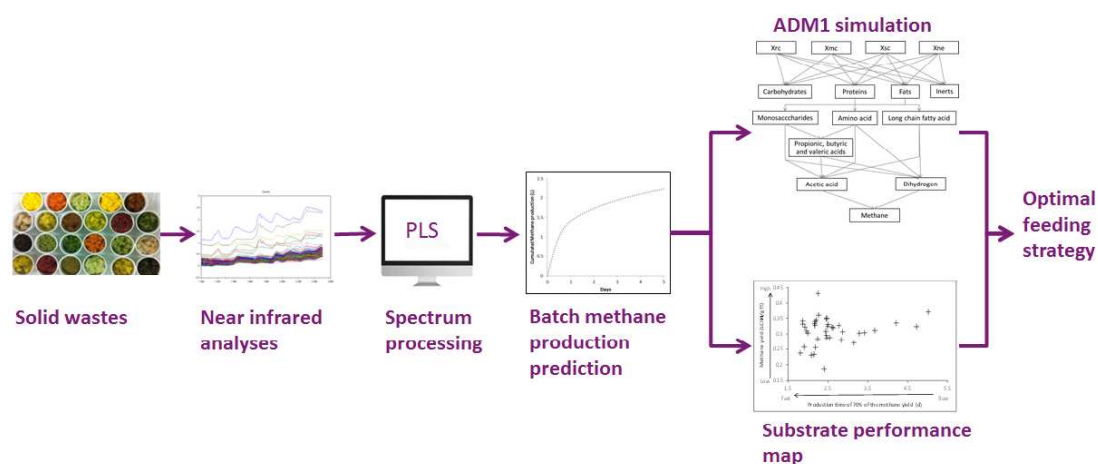
Accepted Date: 7 May 2017

Please cite this article as: Charnier, C., Latrille, E., Jimenez, J., Torrijos, M., Sousbie, P., Miroux, J., Steyer, J.-P., Fast ADM1 implementation for the optimization of feeding strategy using near infrared spectroscopy, *Water Research* (2017), doi: 10.1016/j.watres.2017.05.051.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Comment citer ce document :

Charnier, C. (Auteur de correspondance), Latrille, E., Jimenez, J., Torrijos, M., Sousbie, P., Miroux, J., Steyer, J.-P. (2017). Fast ADM1 implementation for the optimization of feeding strategy using near infrared spectroscopy. *Water Research*, 122, 27-35. DOI : 10.1016/j.watres.2017.05.051



# Fast ADM1 implementation for the optimization of feeding strategy using near infrared spectroscopy

Cyrille Charnier <sup>a b</sup> (cyrille.charnier@bioentech.eu), Eric Latrille <sup>a</sup> (eric.latrille@inra.fr), Julie Jimenez <sup>a</sup> (julie.jimenez@inra.fr), Michel Torrijos <sup>a</sup> (michel.torrijos@inra.fr), Philippe Sousbie <sup>a</sup> (philippe.sousbie@inra.fr), Jérémie Miroux <sup>b</sup> (jeremie.miroux@bioentech.eu), Jean-Philippe Steyer <sup>a</sup> (jean-philippe.steyer@inra.fr).

<sup>a</sup> LBE, INRA, 102 Av. des Etangs, F-11100 Narbonne, France

<sup>b</sup> BioEnTech, 74 Av. Paul Sabatier, F-11100 Narbonne, France

Corresponding author: Cyrille Charnier (E-mail: [cyrille.charnier@bioentech.eu](mailto:cyrille.charnier@bioentech.eu))

## Abstract

Optimization of feeding strategy is an essential issue of anaerobic co-digestion that can be greatly assisted with simulation tools such as the Anaerobic Digestion Model 1. Using this model, a set of parameters, such as the biochemical composition of the waste to be digested, its methane production yield and kinetics, has to be defined for each new substrate. In the recent years, near infrared analyses have been reported as a fast and accurate solution for the estimation of methane production yield and biochemical composition. However, the estimation of methane production kinetics requires time-consuming analysis. Here, a partial least square regression model was developed for a fast and efficient estimation of methane production kinetics using near infrared spectroscopy on 275 bio-waste samples. The development of this characterization reduces the time of analysis from 30 days to a matter of minutes. Then, biochemical composition

and methane production yield and kinetics predicted by near infrared spectroscopy were implemented in a modified Anaerobic Digestion Model n°1 in order to simulate the performance of anaerobic digestion processes. This approach was validated using different data sets and was demonstrated to provide a powerful predictive tool for advanced control of anaerobic digestion plants and feeding strategy optimization.

#### **Keywords**

Spectroscopy; Kinetics; Methane potential; Modeling; Anaerobic digestion; ADM1.

28

29 **1 INTRODUCTION**

30

31 Anaerobic digestion (AD) is probably the major biological process involved in waste valorization. It  
32 has a low energy requirement, low sludge production and large potentialities to produce both valuable  
33 intermediates and energy. Nonetheless, diversification of substrates and development of Anaerobic co-  
34 Digestion (AcoD) are needed in order to keep AD as a highly efficient process. Since feeding strategy is  
35 the major factor governing the overall process, the development of AcoD involves optimal feeding  
36 strategy definition (Karthikeyan and Visvanathan, 2013). The variability of the available inputs and the  
37 diversity of technological processes has made the forecasting of the process performances and the  
38 definition of an optimal feeding strategy difficult. In particular, precise substrates characterization and  
39 advanced dynamic models are required to predict the performances of the process depending on its  
40 feeding strategy (García-Gen et al., 2013). In that aim, the Anaerobic Digestion Model #1 (ADM1) has  
41 been developed (Batstone et al., 2002). ADM1 is a detailed model taking into account the main reactions  
42 of AD processes. ADM1 has been extensively used and analyzed in both academic and practical  
43 applications (Batstone et al., 2006).

44 The first limit is that hydrolysis is recognized as the rate-limiting step in the complex digestion  
45 process (Appels et al., 2008) but is represented as a single-order reaction in ADM1. Contois-type  
46 equation represents more accurately the hydrolysis of solid substrates (Yasui et al., 2006). Moreover,  
47 García-Gen et al., (2015) and Jimenez et al., (2016) recommended to split the organic matter into three  
48 to four compartments with different kinetics of hydrolysis to fit the experimental biogas production. The  
49 second limit encountered while using such a complex model is that many parameters have to be

determined and depend either on the substrate or on the process (Batstone et al., 2015). To set these parameters, a precise characterization of the substrate composition as input variables is needed. Input variables of the model describe the substrate in terms of protein, lipid, carbohydrate content and Chemical Oxygen Demand (COD), biodegradability and kinetics of hydrolysis. Thus, several methods have been proposed to simplify the implementation. Basically, Folin reagent or Kjeldahl procedure are used to estimate the protein content (Souza et al., 2013), hexane extraction for lipid content (Girault et al., 2012) and colorimetric assay using a complexation between phenol and glucose for the carbohydrate content estimation (Ugwuanyi et al., 2005). Kleerebezem and Van Loosdrecht, (2006) proposed to use the elemental composition of the substrate and to transform it into proteins, lipids, carbohydrates, and COD. Zaher et al., (2009) generalized this approach creating a transformer that converts the elemental composition and COD into an input ADM1 vector. The limit of the approaches based on elemental composition is the estimation of the biodegradability and kinetics of hydrolysis. Biodegradability estimation is usually calculated as a rate between COD and Biochemical Methane Potential (BMP). BMP is experimentally obtained by batch digestion experiments which are time-consuming (Angelidaki et al., 2009). However, the laboratory analyses reported above are time consuming and labor intensive, which is an important drawback for industrial application. These characterizations of the substrate are rarely available at industrial scale and must be determined for each new substrate (Appels et al., 2011). The difficulty to estimate the parameters describing the substrate reduces the applicability of the model for the predictions of the digestion performances at industrial scale.

Thus, different methods have been proposed to shorten the time of analysis. Jimenez et al., (2014) used a Partial Least Square (PLS) regression model to predict biodegradability by 3D fluorescence coupled with chemical sequential extractions. It allows reducing the time of analysis from one month to one week. Another method has been proposed by Lesteur et al., (2011) and Jacobi et al., (2012) which reduces the time of BMP analysis to a matter of minute using Near InfraRed (NIR) spectra. NIR spectra

are mathematically treated using Standard Normal Variate normalization (SNV), detrending or Savitsky-Golay filtering and used to predict the BMP thanks to a PLS model (see also Ward, 2016). Recently, Charnier et al., (2016) extended the use of the NIR analyses to the prediction of the carbohydrate, protein and lipid content and chemical oxygen demand. Combining the literature on the prediction of the BMP and the biochemical composition, NIR appears to be an efficient method for a fast implementation of the ADM1 model.

Kinetic estimation remains a challenge; few methods are available and are all based on biological digestion of the substrates (Jensen et al., 2011). From these digestion experiments, the methane production rate (MPR) is obtained and the kinetic is estimated so that the model fits the MPR (Yasui et al., 2008). Batstone et al., (2009) suggested that this approach underestimates the kinetics of methane production because kinetics depends on the micro-organisms adaptation to the substrate. In order to adapt the micro-organisms to the substrate, García-Gen et al. (2015) estimated the kinetics on eight successive batches using the same substrate to feed the reactor. Thus, methane production kinetics estimation requires batch or continuous digestion experiments which make the kinetics of methane production one of the longest parameter to obtain. To shorten the estimation of methane production kinetics is a hindrance to overcome in order to use ADM1 for industrial purposes (Batstone et al., 2015).

Hence, this study aims at reducing the difficulties of ADM1 parameters estimation. It first focuses on the possibility to predict the methane production kinetics of solid substrates using NIR spectroscopy. In a second time, combining the models existing on the literature (Charnier et al., 2016; Jacobi et al., 2012; Lesteur et al., 2011) and the model developed in this study, NIR is used to predict biochemical composition and MPR to implement ADM1. The hydrolysis kinetics of ADM1 is optimized to fit the MPR predicted by NIR. This approach reduces the ADM1 parameters estimation such as BMP or hydrolysis



kinetic from 30 days to a matter of hours, making it suitable for industrial application. Eventually, the accuracy of this approach has been estimated based on two case studies.

## **2 MATERIAL AND METHODS**

### **2.1 Batch experiments**

MPR was estimated from biological degradation of the substrate. Experiments were carried out in 8.5 L batch reactors with 6 L of sludge under mesophilic conditions as described in García-Gen et al., (2015). The substrates were grinded to approximately 1cm. In order to adapt the micro-organisms to the substrate, The reactor was fed at  $1 \text{ gVS.l}^{-1}$  with the same substrate in eight successive batches, each one lasting for only few days, due to the low amount of substrate added. The biogas produced passed through a moisture trap and through a milligas counter fitted with a digital output (MGC-1 gas flow meters, DR.-ING.RITTER APPARTEBAU GMBH & CO.KG, Bochum, Germany); the data were recorded on-line. The methane production of the last batch (i.e., after adaptation of the microbial ecosystem to the substrate) was used to obtain the MPR. Methane yield was estimated as the maximal methane produced from 1 gTS on the eighth batch. Methane production times were selected on the MPR as values corresponding to the durations required to reach a certain ratio of the methane yield from 5 to 95 % with a regular increment of 5 %.

### **2.2 NIR analyses**

According to Lesteur et al., (2011), freeze-dried and grinded samples were scanned in reflectance over 1000-2500 nm, with a resolution of  $2.5 \times 10^6 \text{ nm}$ , using a spectrophotometer BUCHI NIR-Flex N-500 solids fitted with vial accessory (Buchi®, Flawil, Switzerland). For the first chemometric treatment, absorbance (A) was calculated from reflectance (R) as  $A = \log(1/R)$  and SNV was applied. For the second

treatment, Savitzky–Golay filter with a second derivative was added to previous treatment of the spectra. Eventually, a third treatment was added, based on the first treatment but with a detrending as an additional chemometric treatment.

### 2.3 Samples analyzed

275 samples representing a wide diversity of substrates were analyzed to build the PLS model for the prediction of the methane production times and BMP. In details, the analyses were conducted on 3 fat or oleaginous samples, 24 fruits, 18 vegetables, 12 farm wastes such as straw, manure or grass, 15 cereals, 6 meat or fish extracts, 3 micro-algae and 194 mixed organic wastes. Based on the recommendation of Dardenne (2010), the dataset was split in an independent validation set and a calibration set before establishing any correlation with NIR spectroscopy. A validation set of 42 samples was separated for the model validation composed of 6 vegetables and 36 mixed organic wastes. The remaining 237 samples were used to calibrate the PLS regression model. All the samples were freeze-dried and grinded below 1 mm before any analysis. The mean methane yield was about  $0.316 \text{ L.gTS}^{-1}$  and  $0.270 \text{ L.gTS}^{-1}$  with a standard deviation of  $0.093 \text{ L.gTS}^{-1}$  and  $0.041 \text{ L.gTS}^{-1}$ , a minimal value of  $0.129 \text{ L.gTS}^{-1}$  and  $0.167 \text{ L.gTS}^{-1}$ , a maximal value of  $0.717 \text{ L.gTS}^{-1}$  and  $0.351 \text{ L.gTS}^{-1}$  for the calibration and validation sets respectively. Figure 1 represents the times required to reach a certain ratio of the methane yield, from 5 to 95%, of both calibration and validation sets. The same samples as described in (Charnier et al., 2016) were used to build the PLS model for the prediction of the biochemical composition.

### 2.4 Partial least square regression for kinetics prediction.

Methane yield was predicted using the toolbox FACT on Scilab 5.5.0. Methane production times were firstly considered as independent values when running the PLS regression with the tool box FACT, Scilab® 5.5.0. Each of the discrete times (from 5% to 95%) was used to build a specific PLS model leading to 19 independent models with no dependencies between the latent variables for the 19 models.

In a second approach, the discrete times (from 5% to 95%) were considered as a vector using the package MixOmics, R<sup>®</sup> 3.2.2 leading to one PLS model with a single set of latent variables. In this case the PLS model was built using MixOmics.

In both approach, the PLS models for the prediction of the methane production times and the PLS model for the prediction of the methane yield were calibrated using the data from the calibration set. Cross-validations were performed on the calibration set using the Jackknife method (Karaman et al., 2013) which splits randomly the dataset into 15 independent blocks. For each PLS model, the performances of the cross-validation were used to determine the number of latent variables and to choose the best chemometric treatment. The final model was then applied to the independent validation set. The accuracy and robustness of the model were assessed with the indicators recommended in Dardenne (2010). Cross-validation performances were analyzed based on the model Standard Error of Calibration (SEC), the Standard Error of Cross Validation (SECV) and the coefficient of determination between predicted and reference values ( $R^2$ ). The SEC was unfortunately not accessible using MixOmics but the  $Q^2$  and SECV were used to assess the performance of the PLS. The accuracy of the PLS on the validation set was assessed by the Standard error of Prediction (SEP) and bias.

## 2.5 Modified ADM1

ADM1 benchmarked on the BSM2 report (Rosén and Jeppsson, 2008) was implemented in Scilab 5.5.0. ADM1 model was then modified based on the work of (Ramirez et al., 2009). The degradation of the particulate organic matter has been implemented with Contois kinetics to describe both disintegration and hydrolysis. Hydrolysis was supposed to be non-limiting. Three micro-organism groups were considered for the hydrolysis of carbohydrates, proteins and lipids. A consortium regrouping all the hydrolytic micro-organism was considered for the disintegration. An additional modification of this

model was tested, as recommended in (Jimenez et al., 2016), four different particulate inputs were implemented, Xrc, Xmc, Xsc and Xne according to their bioaccessibility (Figure 2). Xrc, Xmc, Xsc and Xne stand respectively for the readily accessible, moderately accessible, slowly accessible and not extractible organic matter. These four particulate inputs are considered to have the same composition and to be degraded by the same biomass but with different  $\mu_{\max}$  in this modified ADM1.

## 2.6 MPR prediction

The prediction of the MPR curve is made using the predicted methane yield and the 19 methane production times (5-95 %) predicted. At each of the methane production time predicted is associated a methane production using the methane yield multiplied by the percentage reached. Thus 19 points of the MPR curve are estimated based on the NIR prediction. The initial point was added as zero methane produced at time zero. The MPR was then interpolated. A smooth spline with the function *smooth.spline* was done using R<sup>®</sup> 3.2.2 to remove noise. For the prediction of the confidence range,  $\pm 1$  SEP was added to the 19 methane production times and the methane production yield before to produce the MPR using this approach, in order to transfer the potential errors of the NIR prediction on the predicted MPR.

## 2.7 Biochemical composition prediction

The same methods, datasets and models as reported in (Charnier et al., 2016) were used for an accurate prediction of the carbohydrates, nitrogen and lipids content and COD. Thus, the same performances of prediction as described in (Charnier et al., 2016) were obtained. The total carbohydrates content was predicted in  $\text{mgO}_2.\text{gTS}^{-1}$ , the total nitrogen content was predicted in  $\text{g.gTS}^{-1}$ , the total lipids content was predicted in  $\text{g.gTS}^{-1}$  and the total COD was predicted in  $\text{mgO}_2.\text{gTS}^{-1}$ .

## 2.8 ADM1 modelling approach

Conversion of the NIR predicted values into ADM1 parameters are presented in Figure 3. Total nitrogen is converted into protein COD equivalent multiplying the total nitrogen in gram by  $6.25 \times 1.54$  (Girault et al., 2012).  $\frac{1}{6.25}$  is the nitrogen content in protein and 1.54 the mean COD values of proteins in  $\text{gO}_2 \cdot \text{g}^{-1}$ . Carbohydrates content was directly estimated in COD equivalent. Lipids content was estimated in gram per gram, thus a conversion rate of 2.8 between gram and COD was applied (Batstone et al., 2002). 2.8 represents the averaged COD of 1 gram of oil in  $\text{gO}_2 \cdot \text{g}^{-1}$ . COD and methane yield predictions were used to assess the non-biodegradable content of the substrate using the equation 1 described by Jimenez et al., (2014):

$$\text{Biodegradability} = \frac{\text{BMP}}{(\text{COD} \cdot 350)}$$

(1)

Biodegradability is expressed as a ratio from 0 to 1, BMP in  $\text{NmLCH}_4 \cdot \text{g}^{-1}$  and COD in  $\text{gO}_2 \cdot \text{g}^{-1}$ . If  $X_{rc}$ ,  $X_{mc}$ ,  $X_{sc}$  and  $X_{ne}$  inputs were implemented, the repartition of the initial COD in  $X_{rc}$ ,  $X_{mc}$ ,  $X_{sc}$  and  $X_{ne}$  was determined by analyzing the changes in the methane production slope of the predicted MPR. For the determination of the initial state of the model variables, a maximal specific growth rate ( $\mu_{\max}$ ) of  $3 \text{ d}^{-1}$  was set, representing the mean  $\mu_{\max}$  found in the 275 samples. The initial state of the ADM1 variables was simulated in accordance with the experimental protocol (see Material and Methods section 2.1). Seven batches were successively simulated with an input of  $1 \text{ gTS} \cdot \text{l}^{-1}$  in a 6 liter reactor under mesophilic conditions. The values of the initial state of ADM1 variables were taken as the final values reached on the seventh batch. Then, on the eighth simulated batch, the kinetics parameters of ADM1 were assessed based on the MPR predicted by NIR. The estimation of the kinetics was considered as a non-linear problem based on the root mean square optimization between the MPR predicted by NIR and the MPR simulated by ADM1. Using the modified ADM1 (see Material and Methods 2.5), the  $\mu_{\max}$  of the Contois kinetic terms were optimized by the operator in order to minimize the square value of the difference

between the predicted methane production and the simulated methane production. In order to assess the improvements brought by the modifications of the model, the  $\mu_{\max}$  of the first order disintegration kinetic of the standard ADM1 implementation obtained from (Rosén and Jeppsson, 2008) was also optimized. This was done minimizing the square value of the difference between the predicted methane production and the simulated methane production using the function *leastsq* from Scilab<sup>®</sup> 5.5.0.

### 3 RESULTS AND DISCUSSION

#### 3.1 Prediction of the methane production rate in batch conditions using NIR spectroscopy

##### 3.1.1 Prediction of the methane yield

The methane yield was predicted according to (Lesteur et al., 2011) using the same chemometric treatment, presented in Material and Methods section **Error! Reference source not found.** (absorbance, SNV and Savitzky–Golay filter with a second derivative). Using the data set of the present study and based on the cross-validation performances, the PLS regression model was calibrated using 13 latent variables. The number of latent variables was selected to minimize the SECV. On the calibration set, the SEC was about 0.019 L.gTS<sup>-1</sup>, SECV about 0.026 L.gTS<sup>-1</sup> and R<sup>2</sup> about 0.92. The model was validated on the independent validation set with SEP of 0.025 L.gTS<sup>-1</sup>. The SEP found for the methane yield prediction is comparable with the SEP of 0.028 L.gVS<sup>-1</sup> reported in (Lesteur et al., 2011). It validates the use of NIR infrared to predict methane yield in this study. If previous models give predictions per gram of volatile solids, predictions per gram of total solid have been preferred since the total solid is faster to estimate than volatile solid and prediction performances were similar.

### 3.1.2 Prediction of the methane production times

First of all, the times required to reach a certain percentage of the methane yield ranging from 5 to 95 % of the final value has been considered as independent values at each ratio of the methane yield. The choice of the chemometric treatment was based on the PLS performance during cross-validation and calibration. SEC, SECV and  $R^2$  between predicted and reference values in cross-validation are represented in Figure 4. For all the treatments, SEC and SECV increase from 5 to 70% of the methane yield, this corresponds to an approximate 6-fold change. SEC and SECV then decrease with the methane yield ratio from 70 % to 95 %. The variation of SEC and SECV is correlated with the variance within the dataset which increases with the ratio of the methane yield to reach a maximal variance between 65 % and 70 % of the methane yield as suggested in Figure 1. The experimental standard deviation and the standard deviation of the predicted values in cross-validation are plotted in Figure 4. The standard deviation reaches a maximum of 1.24 day at 65 % of the methane yield. Thus, the most important difference of production time between the samples is reached at 65 % of the methane yield. It explains the increase of both SEC and SECV up to this point. Then, the variance decreases indicating a homogenization of the samples. It can be concluded that 65 % of the methane yield is the most discriminant production time. It can be observed in figure 4 that, regardless of the chemometric treatment, using independent PLS model,  $R^2$  increases with the methane yield ratio. It means that model predictions are less accurate on early methane production. It can be mainly explained by two factors. First, to restore anaerobic conditions after having fed the digester, the head space is flushed using nitrogen. As a consequence, the very first biogas production is due to both liquid gas equilibrium and substrate biodegradation, which disturbs the estimation of the methane produced from the substrate biodegradation. Secondly, the early methane production is due to 5 to 15 % of the organic matter. Thus the information related to this early production can be diluted in the NIR spectrum which is made on the overall substrate.

About the chemometric treatments, both SEC and SECV were minimized using treatment absorbance, SNV and Savitzky–Golay filter with a second derivative at any ratio of the methane yield with respectively values from 0.05 to 0.37 days and 0.07 to 0.55 days.  $R^2$  was maximal using treatment absorbance, SNV and Savitzky–Golay filter with a second derivative until 70% of the methane yield whereas treatment absorbance, SNV and detrending slightly improved  $R^2$  after 70% of the methane yield compared to treatment absorbance, SNV and Savitzky–Golay filter with a second derivative. Treatment absorbance, SNV and Savitzky–Golay filter with a second derivative maximized the PLS regression performances for the prediction of production time and was chosen as the reference chemometric treatment for this purpose.

PLS regression was also built to treat the production time as a vector of correlated values. In this case, a global set of latent variables is used. The advantage of this approach is to acknowledge the correlation between the production times at any ratio of the methane yield. The disadvantage is the necessity to set a unique number of latent variables for every ratio of the methane yield. Using treatment absorbance, SNV and Savitzky–Golay filter with a second derivative, both approaches were compared considering either the time of production as a correlated vector or as independent values. SEC was not accessible using MixOmics and is thus absent of the Figure 4.  $Q^2$  was used to set the number of latent variables to 5 which corresponds to the first  $Q^2$  value below 0.0975 with a local minimum. The number of latent variables considering production time as independent values varies from 6 to 12. A reduction of the number of latent variables is usually acknowledged as a gain in robustness. Nonetheless, SECV was higher considering production time as a vector with values from 0.07 to 0.60 days than considering the methane production times as independent values. Moreover, the  $R^2$  was lower with values from 0.47 to 0.79. On average, considering the methane production time as a vector increased the SECV by 8.5% and decreased  $R^2$  by 6.1%. From these results, it can be concluded that the



performances are maximized considering the time of production as independent values with treatment absorbance, SNV and Savitzky–Golay filter with a second derivative.

The resulting PLS regression model was then tested on the validation set. SEP values are presented in Figure 5. The same patterns as for SECV and SEC estimation were observed for SEP and RMSEP which reached maximal values at 0.65 point of the methane yield with respectively 0.47 and 0.46 day. The standard deviation of the validation set confirmed an increase of the dispersion until 65 % of the methane yield (Figure 5). SEP and RMSEP are approximately 2 times lower than the standard deviation of the validation set which proves that PLS was able to discriminate the production time of the samples. It also shows that the errors of prediction are low, for instance at 65 % of the methane yield the SEP was 0.47 day to predict values in a range from 1.15 to 4.75 day with a mean value of 2.32 day. The model previously built on the calibration set can thus be considered as suitable to predict the times of degradation at any methane yield ratio.

The cross-validation and validation results using treatment absorbance, SNV and Savitzky–Golay filter with a second derivative and considering the times required to reach a certain percentage of the methane yield as independent values are detailed in the bottom of Figure 5. In these two figures, the methane production times are represented without any distinction on the methane yield ratio. Figure 5 confirms the accuracy of the PLS model but also shows that the error of prediction decreases with the predicted time. The methane production times above 6 days mostly represent a methane production above 65 % of the final methane yield. It has been previously observed that the variance decreased after 65 % of the methane yield and that this goes with a decrease of SECV, SEP and RMSEP, which explains the decrease in the error of prediction. From the cross-validation results, it can be seen that one sample is strongly under-predicted at several methane yield ratios. It is actually the only sample represented by oil. It shows that oil is under-estimated on the current dataset of calibration. This error is not found

during the validation step because no oil was contained in the validation set. Nonetheless, from these results, it can be concluded that the prediction of production time is feasible using NIR spectroscopy and PLS model, even if an improvement of the dataset could be made including more oil samples.

### **3.1.3 Prediction of the methane production rates**

A case study was carried out on two samples from the validation set to assess the feasibility of the MPR prediction. As explained in the section 2.6, the MPR was predicted using both methane production times and methane yield predicted by NIR analyses. Then MPR was interpolated from these 19 points. These curves are represented as “NIR predictions” in Figure 6 and compared to the experimental cumulated methane production. It shows that NIR prediction represents properly the production of methane. Both methane yield and production time were correctly predicted. The errors of prediction on the methane yield were about 6.7 % on salad and 3.5 % on cauliflower. On the experimental data, the methane flow decreased after 0.4 and 1 day and increased again after 1 and 1.5 day for respectively salad and cauliflower (Figure 6.A-B). These behaviors were not accurately predicted by NIR. One possible explanation is an inhibition of the readily degradable substrate on the slowly degradable substrate that was not predicted by NIR. Another hypothesis leading to similar results can be found in Yasui et al., (2008): the readily accessible organic matter is degraded first followed by the slowly accessible organic matter, keeping in mind that kinetics of disintegration and accessibility are independent.

### **3.1.4 Simple predicted indicator of methane production performances**

MPR is a relevant indication that allows to differentiate the methane production performances of two substrates, but MPR analysis becomes difficult when the number of samples to compare increases. Methane yield and kinetics of methane production are two key parameters to elaborate an optimal feeding strategy. It is indeed in the interest of the plant operator to maximize both methane yield and kinetics. The time needed to reach 70 % of the methane yield has been reported in section

3.1.2 to be the most discriminant time. The observation of this time allows differentiating a substrate with a fast kinetic of methane production. In Figure 7, the samples of the validation set have been plotted depending on their predicted methane production time needed to reach 70% of the methane yield and their predicted methane yield. Figure 7 shows that the fast NIR spectroscopy analysis of the substrate provides the opportunity to the plant operator to select the most relevant substrates on a simple graph.

### 3.2 ADM1 calibration from NIR prediction

It has been proven that NIR spectroscopy can be used to estimate MPR in batch conditions. It turns to be also a powerful method to implement an ADM1 model able to predict the substrate methane production performances. The approach was applied to the case studies of salad and cauliflower. As described in Materials and Methods section **Error! Reference source not found.**, carbohydrates, nitrogen and lipids content, COD and methane production rate were used to implement ADM1. The curve predicted by NIR was then used as a reference for the  $\mu_{\max}$  optimization of the Contois kinetic parameters. The model using only one particulate input (Batstone et al., 2002) was first tested. Respectively for salad (Figure 6.A) and cauliflower (Figure 6.B),  $\mu_{\max}$  were estimated at 1.92 and 4.33 d<sup>-1</sup>. It can be seen that, while optimizing  $\mu_{\max}$ , the model does not represent accurately the predicted MPR. Indeed, it tends to under-estimate the kinetics at the beginning of the production and to over-estimate it in the second part. More kinetic parameters are thus needed to represent efficiently the curve. Hence, a second optimization was done using the modified ADM1 described in section 2.5 with four fractions and four distinct  $\mu_{\max}$  kinetic parameters for disintegration. It can be seen that four kinetics of disintegration improved the fit between the simulated cumulated biogas production and the predicted cumulated biogas production. The degradation of X<sub>rc</sub> is faster than when considering a single input while the degradation of X<sub>mc</sub>, X<sub>sc</sub> and X<sub>ne</sub> are slower, which fits the predicted data. It can be seen in Figure 6.A-B

that the NIR predicted data allows calibrating a modified ADM1 that accurately represents the experimental batch data.

If we consider a normal distribution of errors, 68.2 % of the reference values are contained in a range of  $\pm 1$  SEP of the predicted value. Thus, an error of  $\pm 1$  SEP was applied to each methane production time, to the methane yield and to the parameters describing the biochemical composition predicted by NIR in order to obtain an envelope where 68 % of the possible MPR curves will be included.  $\mu_{\max}$  of the Contois kinetic parameters were optimized on the most optimistic MPR, minus one SEP on the methane production time and plus one SEP on the methane yield, and on the most pessimistic MPR, plus one SEP on the methane production time and minus one SEP on the methane yield. The confidence range of the simulations is presented in Figure 6.C-D. It can be seen that the experimental MPR is included in the simulated range of confidence of the MPR. It shows that the predicted confidence range represents accurately the experimental MPR. Since the experiments focus on the methane production under non inhibitory conditions, the impact of the errors of prediction on the biochemical composition cannot be observed. This approach is interesting in order to predict the performance of the plant depending on the feeding strategy. The introduction of potential errors of prediction in the model also provided an estimation of the risk that the plant operator is taking (Južnič-Zonta et al., 2012). Thus, it could provide key information to the human operator to optimize safely the production of a real plant.

### ***3.3 Discussion on the MPR prediction using NIR***

Batstone et al., (2009) have demonstrated that hydrolysis kinetic parameters estimated in BMP tests translated to very poor model performance when BMP-estimated values were used to simulate continuous reactors. On the other hand, García-Gen et al., (2015) demonstrated that the calibrated ADM1 kinetic parameters from batch assays that allow micro-organisms to adapt to individual substrates

can be used to model semi-continuous co-digestion operation treating simultaneously 5 wastes. In our case, we used the same experimental protocol as García-Gen et al., (2015) and thus the model parameters can be assumed to be valid for continuous operation of the digester. The main drawback of the batch experiments with adapted micro-organisms is the time needed, from 6 to 12 weeks, which reduces its use. This study also shows that NIR spectrometry can be efficiently used to predict experimental batch MPR on adapted micro-organisms. Kinetic parameters of the ADM1 model calibrated on the predicted batch MPR can be used for continuous co-digestion simulation. This approach gives the opportunity to replace time-consuming biological experiments by fast NIR analysis for feeding strategy optimization. Two drawbacks must be highlighted. First, NIR remains an expensive and unusual device and only laboratories and large plants can afford to own a NIR spectrometer. Second, the analysis can be done on dry samples only. Two disadvantages are linked to drying: volatile compounds cannot be analyzed and on-line application cannot be considered. A further interesting development or improvement of this technology would be to work with wet samples and install an online sensor able to characterize in real time any input coming to the plant. The obstacle is the absorption of water using NIR that disturb the analysis. Mayer (2015) provides for example some hints about the estimation of methane yield on wet samples using NIR.

## 4 Conclusion

The current study shows that:

- The kinetics of methane production was the main hindrance to overcome for a fast implementation of the ADM1 model using NIR. Thus, a PLS model has been developed to estimate properly the kinetics of methane production from a large set of solid substrates using NIR spectroscopy.

- Based on the actual and previous studies on biochemical composition, methane production yield and kinetics prediction using NIR, it is demonstrated that NIR is a powerful tool for the prediction the main parameters needed to optimize AD plants with a single analysis of few minutes on dry samples.
- The predicted parameters using NIR spectroscopy are transformed in ADM1 parameters to implement a modified ADM1. From this virtual plant, it becomes possible to predict the performance of degradation of any input. It gives access to every parameters estimated in the ADM1 model such as the concentration of VFA, nitrogen or the biogas composition.

This study demonstrates that the use of NIR spectroscopy coupled with PLS regression and ADM1 modeling can help the human operators to monitor the plant. This approach can be used as a powerful tool for plant monitoring and feeding strategy optimization.

## References

- Angelidaki, I., Alves, M., Bolzonella, D., Borzacconi, L., Campos, J.L., Guwy, A.J., Kalyuzhnyi, S., Jenicek, P., Van Lier, J.B., 2009. Defining the biomethane potential (BMP) of solid organic wastes and energy crops: A proposed protocol for batch assays. *Water Sci. Technol.* 59, 927–934. doi:10.2166/wst.2009.040
- Appels, L., Baeyens, J., Degreè, J., Dewil, R., 2008. Principles and potential of the anaerobic digestion of waste-activated sludge. *Prog. Energy Combust. Sci.* 34, 755–781. doi:10.1016/j.peccs.2008.06.002
- Appels, L., Lauwers, J., Degreè, J., Helsens, L., Lievens, B., Willems, K., Van Impe, J., Dewil, R., 2011. Anaerobic digestion in global bio-energy production: Potential and research challenges. *Renew.*

- 408 Sustain. Energy Rev. 15, 4295–4301. doi:10.1016/j.rser.2011.07.121
- 409 Batstone, D.J., Keller, J., Angelidaki, I., Kalyuzhnyi, S. V., Pavlostathis, S.G., Rozzi, A., Sanders, W.T.,  
 410 Siegrist, H., Vavilin, V.A., 2002. The IWA Anaerobic Digestion Model No 1 (ADM1). Water Sci.  
 411 Technol. 45, 65–73. doi:10.2166/wst.2008.678
- 412 Batstone, D.J., Keller, J., Steyer, J.P., 2006. A review of ADM1 extensions, applications, and analysis:  
 413 2002–2005. Water Sci. Technol. 54, 1. doi:10.2166/wst.2006.520
- 414 Batstone, D.J., Puyol, D., Flores-Alsina, X., Rodríguez, J., 2015. Mathematical modelling of anaerobic  
 415 digestion processes: applications and future needs. Rev. Environ. Sci. Biotechnol. 14, 595–613.  
 416 doi:10.1007/s11157-015-9376-4
- 417 Batstone, D.J., Tait, S., Starrenburg, D., 2009. Estimation of Hydrolysis Parameters in Full-Scale Anerobic  
 418 Digesters 102, 1513–1520. doi:10.1002/bit.22163
- 419 Charnier, C., Latrille, E., Jimenez, J., Lemoine, M., Boulet, J.-C., Miroux, J., Steyer, J.-P., 2016. Fast  
 420 characterization of solid organic waste content with near infrared spectroscopy in anaerobic  
 421 digestion. Waste Manag. doi:10.1016/j.wasman.2016.10.029
- 422 Dardenne, P., 2010. Some considerations about NIR spectroscopy: Closing speech at NIR-2009. NIR news  
 423 21, 8. doi:10.1255/nirn.1165
- 424 García-Gen, S., Lema, J.M., Rodríguez, J., 2013. Generalised modelling approach for anaerobic co-  
 425 digestion of fermentable substrates. Bioresour. Technol. 147, 525–533.  
 426 doi:10.1016/j.biortech.2013.08.063
- 427 García-Gen, S., Sousbie, P., Rangaraj, G., Lema, J.M., Rodríguez, J., Steyer, J.P., Torrijos, M., 2015. Kinetic  
 428 modelling of anaerobic hydrolysis of solid wastes, including disintegration processes. Waste Manag.  
 429 35, 96–104. doi:10.1016/j.wasman.2014.10.012

- 430 Girault, R., Bridoux, G., Nauleau, F., Poullain, C., Buffet, J., Steyer, J.P., Sadowski, A.G., Béline, F., 2012. A  
 431 waste characterisation procedure for ADM1 implementation based on degradation kinetics. *Water*  
 432 *Res.* 46, 4099–4110. doi:10.1016/j.watres.2012.04.028
- 433 Jacobi, H.F., Ohl, S., Thiessen, E., Hartung, E., 2012. NIRS-aided monitoring and prediction of biogas yields  
 434 from maize silage at a full-scale biogas plant applying lumped kinetics. *Bioresour. Technol.* 103,  
 435 162–172. doi:10.1016/j.biortech.2011.10.012
- 436 Jensen, P.D., Ge, H., Batstone, D.J., 2011. Assessing the role of biochemical methane potential tests in  
 437 determining anaerobic degradability rate and extent. *Water Sci. Technol.* 64, 880–886.  
 438 doi:10.2166/wst.2011.662
- 439 Jimenez, J., Gonidec, E., Cacho Rivero, J.A., Latrille, E., Vedrenne, F., Steyer, J.P., 2014. Prediction of  
 440 anaerobic biodegradability and bioaccessibility of municipal sludge by coupling sequential  
 441 extractions with fluorescence spectroscopy: Towards ADM1 variables characterization. *Water Res.*  
 442 50, 359–372. doi:10.1016/j.watres.2013.10.048
- 443 Južnič-Zonta, Ž., Kocijan, J., Flotats, X., Vrečko, D., 2012. Multi-criteria analyses of wastewater treatment  
 444 bio-processes under an uncertainty and a multiplicity of steady states. *Water Res.* 46, 6121–6131.  
 445 doi:10.1016/j.watres.2012.08.035
- 446 Karaman, I., Martens, H., Erik, K., Knudsen, B., 2013. Comparison of Sparse and Jack-knife partial least  
 447 squares regression methods for variable selection. *Chemom. Intell. Lab. Syst.* 122, 65–77.  
 448 doi:10.1016/j.chemolab.2012.12.005
- 449 Karthikeyan, O.P., Visvanathan, C., 2013. Bio-energy recovery from high-solid organic substrates by dry  
 450 anaerobic bio-conversion processes: A review. *Rev. Environ. Sci. Biotechnol.* 12, 257–284.  
 451 doi:10.1007/s11157-012-9304-9



- 452 Kleerebezem, R., Van Loosdrecht, M.C.M., 2006. Waste characterization for implementation in ADM1.  
453 Water Sci. Technol. 54, 167–174. doi:10.2166/wst.2006.538
- 454 Kouas, M., Charnier, C., Harmand, J., Sousbie, S.S.P., Steyer, J.P., Torrijos, M., 2002. Kinetic modelling of  
455 anaerobic digestion with new fractionation using simple model and ADM1 1. AD14, IWA Word  
456 Congress on Anaerobic Digestion, Viña del Mar, Chile, November 2015
- 457 Lesteur, M., Latrille, E., Maurel, V.B., Roger, J.M., Gonzalez, C., Junqua, G., Steyer, J.P., 2011. First step  
458 towards a fast analytical method for the determination of Biochemical Methane Potential of solid  
459 wastes by near infrared spectroscopy. Bioresour. Technol. 102, 2280–2288.  
460 doi:10.1016/j.biortech.2010.10.044
- 461 Mayer, F., 2015. Biomethane yield of energy crops and prediction of their biochemical methane potential  
462 with near infrared spectroscopy. PhD thesis, Université de Louvain.
- 463 Ramirez, I., Mottet, A., Carrère, H., Délérès, S., Vedrenne, F., Steyer, J.-P., 2009. Modified ADM1  
464 disintegration/hydrolysis structures for modeling batch thermophilic anaerobic digestion of  
465 thermally pretreated waste activated sludge. Water Res. 43, 3479–92.  
466 doi:10.1016/j.watres.2009.05.023
- 467 Rosén, C., Jeppsson, U., 2008. Aspects on ADM1 Implementation within the BSM2 Framework. TEIE 1–  
468 35.
- 469 Souza, T.S.O., Carvajal, A., Donoso-Bravo, A., Peña, M., Fdz-Polanco, F., 2013. ADM1 calibration using  
470 BMP tests for modeling the effect of autohydrolysis pretreatment on the performance of  
471 continuous sludge digesters. Water Res. 47, 3244–54. doi:10.1016/j.watres.2013.03.041
- 472 Ugwuanyi, J.O., Harvey, L.M., McNeil, B., 2005. Effect of digestion temperature and pH on treatment  
473 efficiency and evolution of volatile fatty acids during thermophilic aerobic digestion of model high

- 474 strength agricultural waste. *Bioresour. Technol.* 96, 707–19. doi:10.1016/j.biortech.2004.06.027
- 475 Ward, A.J., 2016. Near-Infrared Spectroscopy for Determination of the Biochemical Methane Potential :  
476 State of the Art 611–619. doi:10.1002/ceat.201500315
- 477 Yasui, H., Goel, R., Li, Y.Y., Noike, T., 2008. Modified ADM1 structure for modelling municipal primary  
478 sludge hydrolysis. *Water Res.* 42, 249–259. doi:10.1016/j.watres.2007.07.004
- 479 Yasui, H., Sugimoto, M., Komatsu, K., Goel, R., Li, Y.Y., Noike, T., 2006. An approach for substrate  
480 mapping between ASM and ADM1 for sludge digestion. *Water Sci. Technol.* 54, 83.  
481 doi:10.2166/wst.2006.529
- 482 Zaher, U., Li, R., Jeppsson, U., Steyer, J.P., Chen, S., 2009. GISCOD: General Integrated Solid Waste Co-  
483 Digestion model. *Water Res.* 43, 2717–2727. doi:10.1016/j.watres.2009.03.018

484

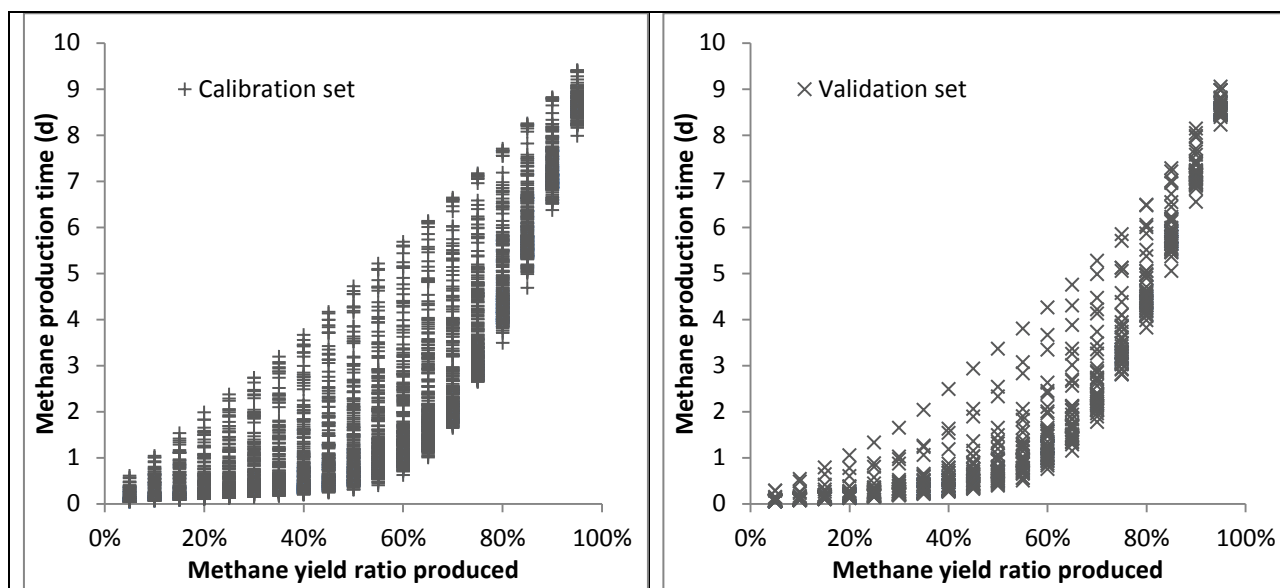


Figure 1: Calibration and validation set features of the experimental data. Description of the methane production time required to reach a certain ratio of the methane yield from 5 to 95 %.

ACCEPTED MANUSCRIPT

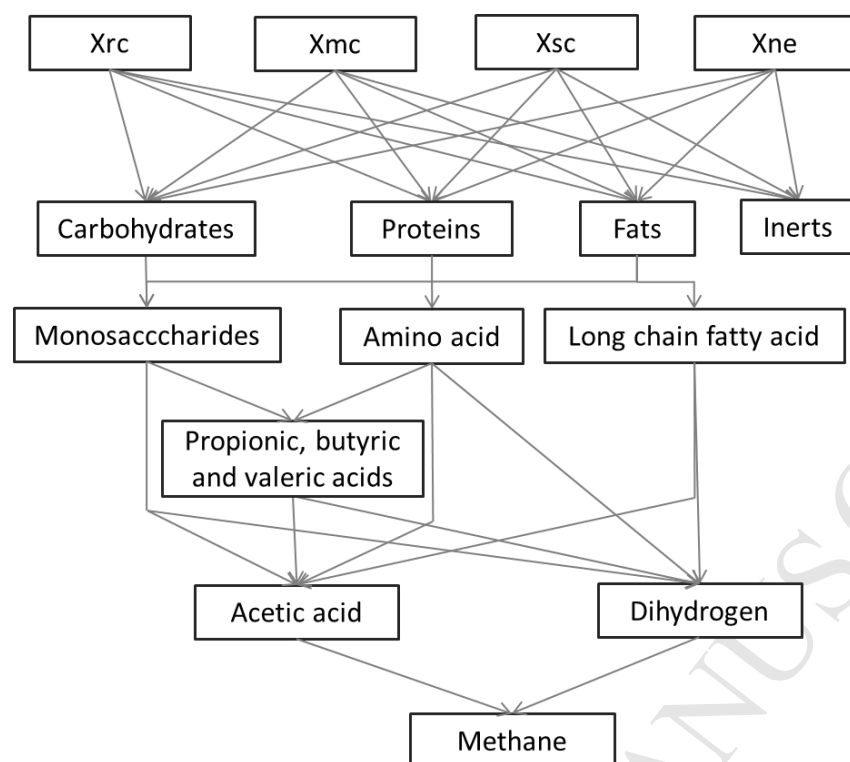


Figure 2: Modified ADM1 with four different particulate inputs Xrc, Xmc, Xsc and Xne.

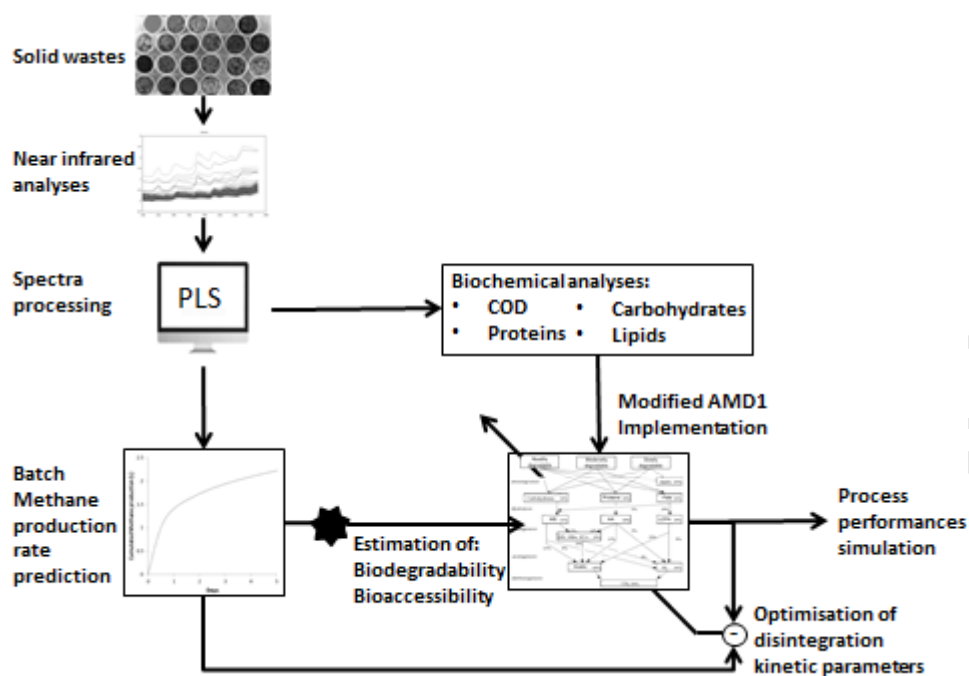


Figure 3: Modeling approach diagram

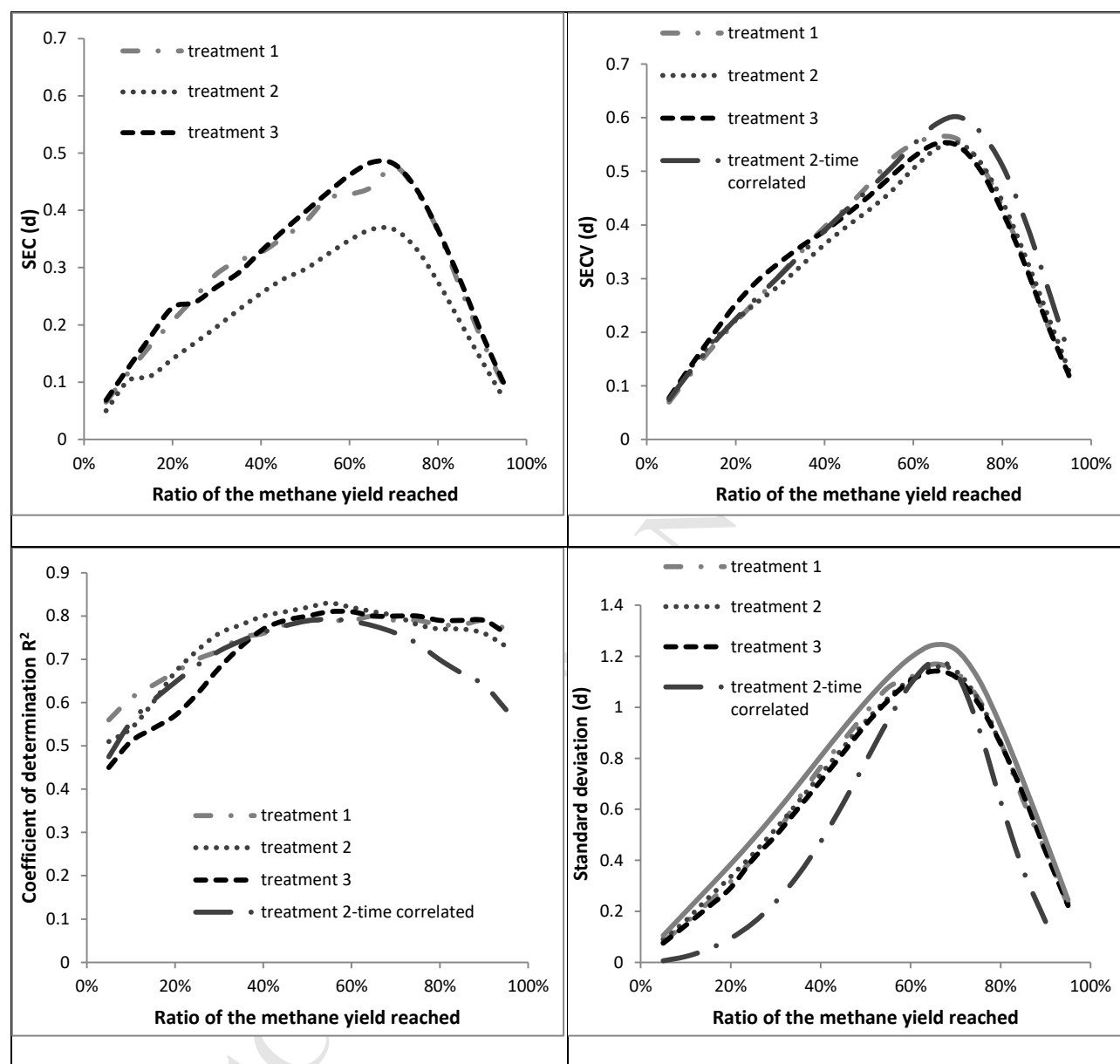


Figure 4: PLS regression performances on cross-validation with 3 different chemometric treatments and considering the methane production times as correlated or independent values. Treatment 1 correspond to absorbance and SNV, treatment 2 to absorbance, SNV and Savitzky-Golay filter and treatment 3 to absorbance, SNV and detrending.

Comment citer ce document :

Charnier, C. (Auteur de correspondance), Latrille, E., Jimenez, J., Torrijos, M., Sousbie, P., Miroux, J., Steyer, J.-P. (2017). Fast ADM1 implementation for the optimization of feeding strategy using near infrared spectroscopy. *Water Research*, 122, 27-35. DOI : 10.1016/j.watres.2017.05.051

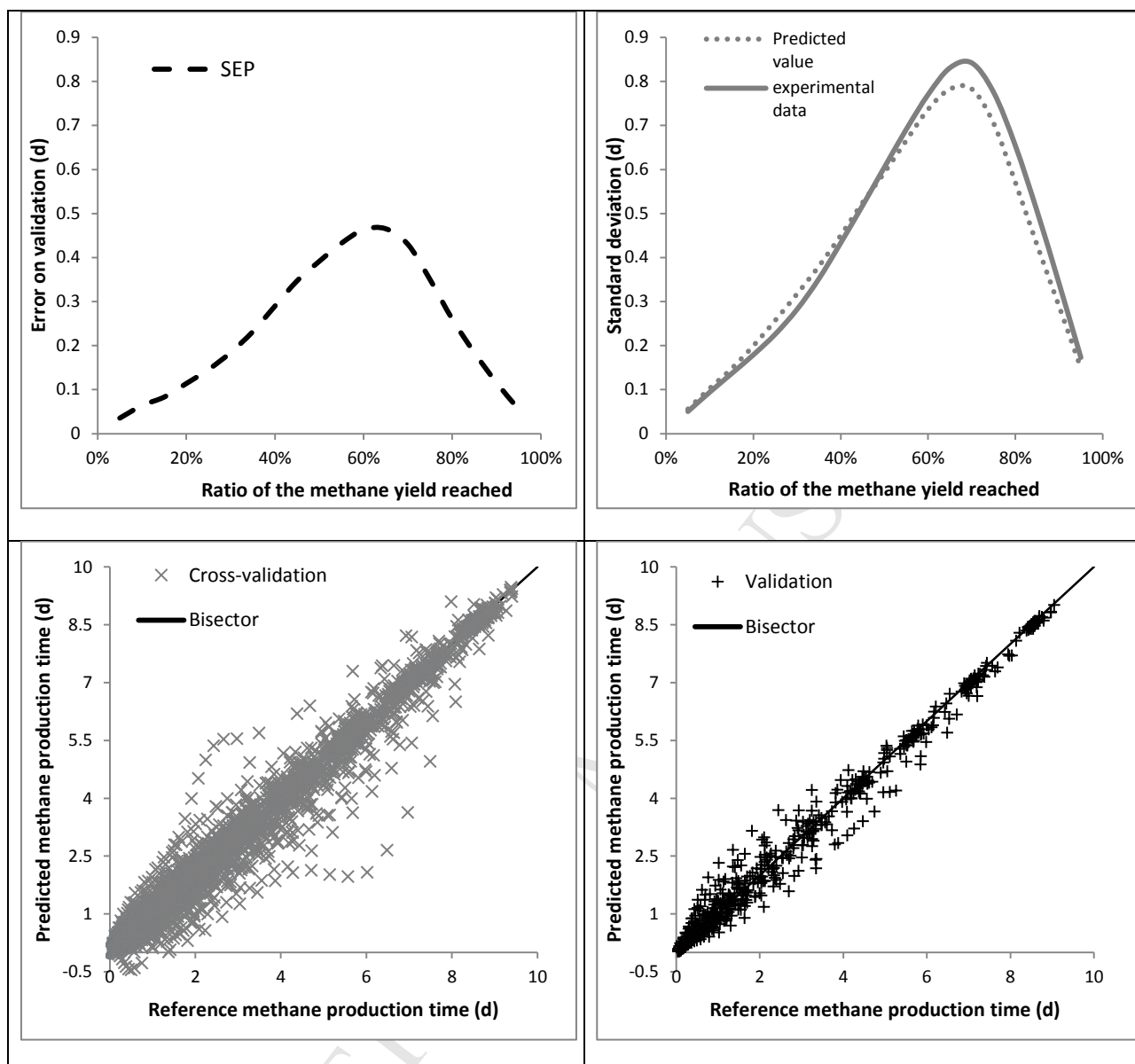


Figure 5: PLS performances on the validation set (up) and representation of the predicted value of both cross-validation and validation (bottom).

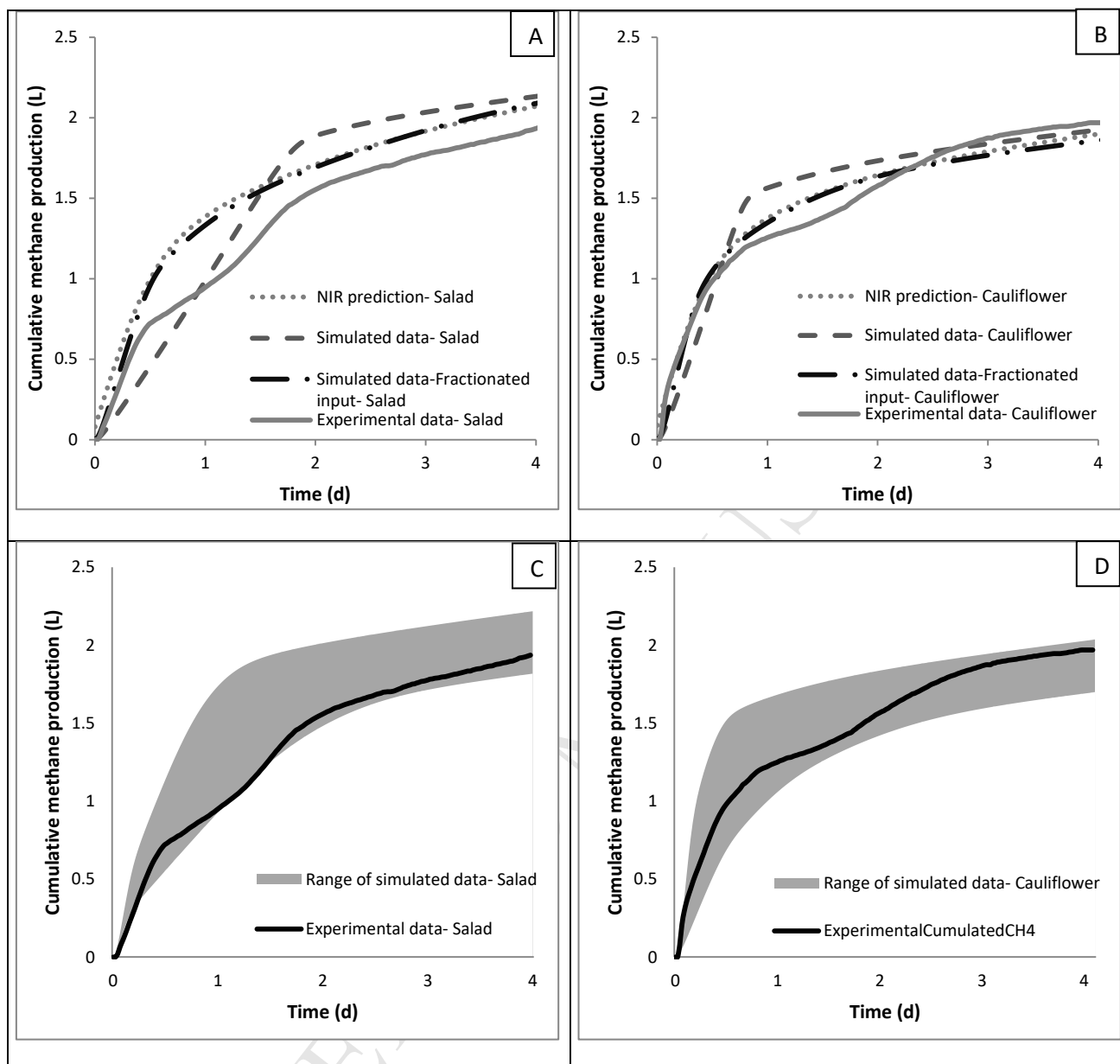
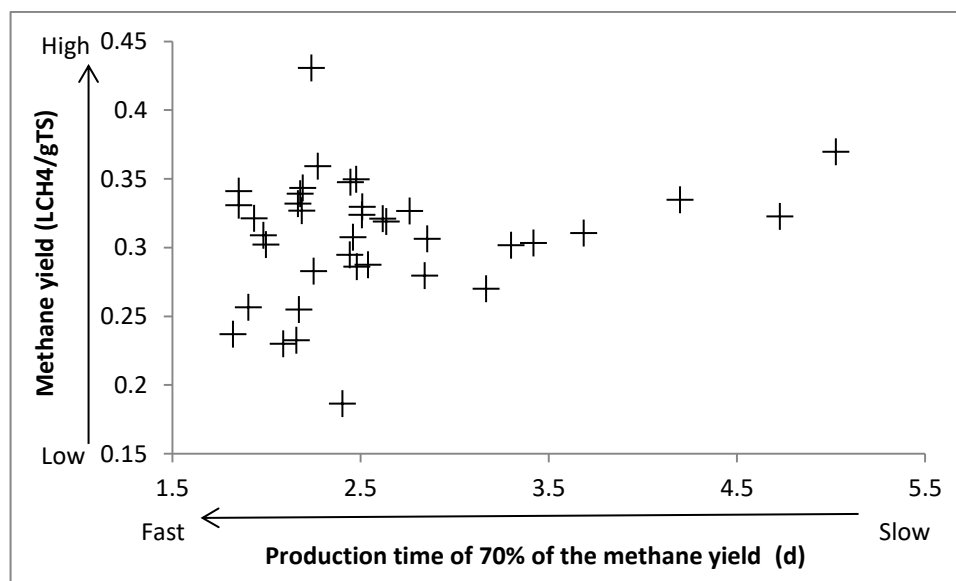


Figure 6: Cumulative methane production curve for anaerobic of salad (left) and cauliflower (right) in batch. A and B, comparison of predicted and simulated data using ADM1 considering the particulate organic matter as a single input or fractioning the particulate organic matter in four inputs with their own kinetics of disintegration for salad and cauliflower respectively. B. C and D, modelling of the methane cumulative production for salad and cauliflower AD taking into account the error of prediction for salad and cauliflower respectively. Methane production times and methane yield were chosen at  $\pm 1$  SEP of the predicted value. The grey area represents the range of confidence.

Comment citer ce document :

Charnier, C. (Auteur de correspondance), Latrille, E., Jimenez, J., Torrijos, M., Sousbie, P., Miroux, J., Steyer, J.-P. (2017). Fast ADM1 implementation for the optimization of feeding strategy using near infrared spectroscopy. *Water Research*, 122, 27-35. DOI : 10.1016/j.watres.2017.05.051





**Figure 7:** Performances of the substrates from the validation set expressed as a function of the methane yield and the production time of 70% of the methane yield. Abscissa arrow indicates the methane production kinetics and ordinate arrow the methane yield.

Comment citer ce document :

Charnier, C. (Auteur de correspondance), Latrille, E., Jimenez, J., Torrijos, M., Sousbie, P., Miroux, J., Steyer, J.-P. (2017). Fast ADM1 implementation for the optimization of feeding strategy using near infrared spectroscopy. *Water Research*, 122, 27-35. DOI : 10.1016/j.watres.2017.05.051

## Highlights

- Near infrared prediction of the methane yield and kinetics
- Use of near infrared spectroscopy to implement ADM1
- Prediction of co-digestion performances assisted with ADM1 simulation
- Feeding strategies optimization using the developed predictive methodology